DOCUMENT RESUME

ED 411 312 TM 027 372

AUTHOR Parkes, Jay
TITLE Performance Assessment and Student Motivation: Questioning Construct Relevant Variance.
PUB DATE 1997-03-00
NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Construct Validity; Correlation; *High School Students; High Schools; *Knowledge Level; Objective Tests; *Performance Based Assessment; Spanish; Student Attitudes; *Student Motivation
IDENTIFIERS *Control of Outcomes; *Variance (Statistical)

ABSTRACT

Whether performance assessments can be claimed to be more inclusive than traditional assessments was studied in an investigation of whether a student's perceptions of control can be detected in a performance assessment score. It was hypothesized that students' perceptions of control would show no effect on their scores on an objective test of knowledge but that they would produce unique variance beyond the knowledge test in the scores on a performance assessment. Seventy-seven 9th graders taking Spanish participated in the study. Perceived control was measured by a scale constructed by J. P. Connell (1985). Students took an objective test and a performance assessment that involved a brief composition in Spanish. Results of a correlation analysis show that the Internal Control scale is significantly though weakly correlated to the performance assessment score. The objective test score is also significantly related to the performance assessment score, although not at the expected level. All study results coincide with the theoretical relationships hypothesized. In this example, the objective test tapped the verb forms and vocabulary students knew, while the performance assessment evaluated what they could use. The question is what the assessor really wants to find out. An appendix presents the performance assessment task. (Contains 1 table and 12 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 411 312

Performance Assessment and Student Motivation: Questioning Construct Relevant Variance

Jay Parkes

The Pennsylvania State University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Jay Parkes

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. March, 1997.

TM027372

BEST COPY AVAILABLE

Performance Assessment and Student Motivation: Questioning Construct Relevant Variance

One of the fundamental concerns with performance assessment is exactly what a performance assessment score represents. Messick (1995) stresses the need to define clearly the construct being assessed so that construct-relevant variance can be distinguished from construct-irrelevant variance. Construct-irrelevant variance is a threat to validity. Determining one from the other is a matter of judgment and is "a tricky and contentious issue (Messick, 1995, p. 743)." It has been claimed for performance assessments that they allow "appropriate room to accommodate students' learning styles, aptitudes, and interests (Wiggins, 1989, p. 712)." This claim raises two questions: first, do these assessments really do that?, and second, if they do, should that be considered construct-relevant or irrelevant? This investigation offers data toward the first question and some speculation about the second.

Claims about performance assessments

Proponents of performance assessments like them, at least in part, because they broaden the scope of information captured by the score from the assessment. Wiggins (1989) holds that individual characteristics about the student may be captured in the score. He has claimed for performance assessments that they allow "appropriate room to accommodate students' learning styles, aptitudes, and interests (Wiggins, 1989, p. 712)." Shepard (1989) calls for assessments which allow conceptual understanding and problem-solving abilities to enter into the score. Also making the claim are Meisels, Dorfman and Steele (1995), who offer a summary of the claims for performance assessment and against standardized assessments. They claim that performance assessments are non-stigmatizing, instructionally relevant, and multi-dimensional; and that they enhance student motivation and promote student learning. Bond (1995) points to the Achilles' heel of such claims: "The list, however, is more in the nature of articles of faith than demonstrated characteristics of the two forms of assessment (p. 21)."

Such demonstrations are just now beginning to appear. Lu and Suen (1995) report that a student's cognitive style interacts with assessment type. That is, field-independent students scored

significantly better on a performance assessment than did field-dependent students, though there was no difference on a multiple choice exam.

The present investigation takes a slightly different approach. Are performance assessments more inclusive, as Wiggins claims, allowing learning styles, for example, to become part of the score? This examination will investigate if a student's perceptions of control can be detected in a performance assessment score.

Perception of control

Connell (1985) discusses three sources to which students attribute control over aspects of their life: an internal control: that they themselves affect what happens to them; a powerful others control: that teachers or other powerful others affect what happens to them; and an unknown control: that the student does not know why things happen to them. Boggiano and Katz (1991) report that those who perceive an internal control are not as susceptible to helplessness behaviors as those with an external control (powerful others or unknown). Helplessness behaviors, according to Boggiano & Katz include "performance impairment, lowered motivation, lowered quality of hypothesis-testing strategies, and reduced pursuit of mastery (p. 36)." They also report that externally-oriented students will prefer easy to challenging tasks and will do the minimum amount of work to achieve the most benefit.

Perceptions of control thus can have direct effects on performance, especially if a task contains elements of choice, creativity, and autonomy. Such a task would favor those with internal perceptions of control. Perceptions of control may also have indirect effects on performance through a correlational network of other variables (Deci & Ryan, 1987).

Internal perceptions of control have been related to intrinsic motivation (Lepper and Hoddell, 1989). Intrinsic motivation, in turn, causes better performance, longer persistence, and challenge-seeking (Csikszentmihalyi and Nakamura, 1989), virtually the mirror image of helplessness. Coming full circle, though, such attributes are only possible if the tasks encourage and allow students to take the control (Lepper and Hoddell, 1989).

Objective tests, by design, allow little room for such concepts to operate. In a multiple choice test, for example, there is no room for creativity, choice or strategy to function. More open-ended tasks *ipso facto* provide room for those functions to occur, but may or may not *require* them. Indeed, the question of whether students *may* use strategies, creativity, etc., versus *must* use them on a performance assessment is one of utmost importance, though it will not be addressed here.

If students' perceptions of control affect the processes they bring to a task when the task permits, differential performance should result between students with internal perceptions and those with external perceptions when those perceptions are able to function. Specifically, in an objective test, no difference should result because the situation does not permit the salient processes to function. In a performance assessment, however, such processes should have room to operate and differential functioning should be observed.

In this investigation, the relationships among a measure of perception of control, an objective knowledge test and a performance measure will be examined. The hypotheses are these:

1. Students' perceptions of control will show no effect on an objective knowledge test.

2. Students' perceptions of control will produce unique variance beyond the knowledge test in the scores on a performance assessment.

Methods

Sample: Seventy-seven ninth grade students in their first year of Spanish participated in the study. The students are from a rural school in Central Pennsylvania. The study was conducted in May, 1994, and the teacher had been giving performance assessments like the one used in the study throughout the year, thus the technique was not new to the students.

Measures: Perceived control was measured by a scale developed by Connell (1985). Only two sections of his scale were used: the cognitive domain and the general domain sections. This measure produces scores in each of the three perceptions of control categories: internal, powerful others, and unknown. Thus each subject receives three scores, one for each perceived control

category summed across the cognitive and general domains. The reliability coefficients for the three scales were: internal control, 0.72; powerful others, 0.53; unknown, 0.81.

The objective test and the performance assessment were designed by the classroom teacher to cover the same material but in a different format. The objective test had 41 completion items that tested knowledge of vocabulary, grammar and idioms covered in the textbook chapter. No reliability data are available for that measure.

For the performance assessment, students were asked to write a paragraph describing a day of sightseeing in Spain. They were allowed to use their textbooks, which included a model paragraph, dictionaries and other reference materials. The assignment is shown in Appendix A. This type of activity required students to synthesize the knowledge of vocabulary, and verbs and tenses acquired from the chapter. The paragraphs were scored by two high school Spanish teachers using a scoring rubric shown in Appendix B. The reliability was assessed through a generalizability study which produced a generalizability coefficient of 0.69.

Procedure: All of the materials and assessments, except the Connell Scale, were part of the students' regular Spanish instruction. The Connell scale was administered while the students were working on the unit but before they completed either assessment. To control for possible practice effects, the four classes of students were counter-balanced for the order in which they completed the assessments, that is, two of the classes did the performance assessment first and two did the objective test first. Analysis of variance revealed no effects due to the order in which the assessments were completed on either assessment.

Analyses: A correlation matrix was calculated first to examine the relationships among the variables. The matrix is given in Table 1. A stepwise regression analysis was then conducted to determine which variables could predict performance assessment score. The results of this analysis are given in Table 2.

Table 1-- Correlation Matrix

	Connell Scales			
	Internal Control	Powerful Others	Unknown	Objective Test
Objective Test	0.17	-0.07	-0.17	
Performance Assessment	0.27*	-0.13	-0.07	0.34*

* $p < 0.05$

Table 2 -- Stepwise Regression Analysis Results

	B	SE B	β
Step 1			
Constant	10.73	2.56	
Obj. Test	0.29	0.093	0.34
$R^2 = 0.11$	F=9.53	$p < 0.05$	
	B	SE B	β
Step 2			
Constant	-2.09	6.66	
Objective Test	0.26	0.092	0.30
Internal Control	0.50	0.24	0.22
$R^2 = 0.16$	$\Delta R^2 = 0.049$	$p < 0.05$	

Results

The results of the correlation analysis show that the Internal Control scale is significantly though weakly correlated to the performance assessment score. The objective test score is also significantly related to the performance assessment score, though, perhaps not at a level one might expect.

The stepwise regression analysis builds on those results and is consistent with them. Only two of the four independent variables were selected via the stepwise procedure to enter the equation: the objective test score and the Internal Control scale score. Both predict unique variance in the performance assessment score.

All of these results coincide with the theoretical relationships hypothesized. First, none of the perceptions of control scales were significantly related to the objective test score. There was no room for them to function. Second, perceptions of control did play a role in the performance assessment score in addition to the "knowledge" the assessment was designed to evaluate as reflected by the objective test. The regression analysis was set up under the hypothesis that the objective test score would reflect the students' cognitive knowledge of the Spanish concepts assessed in both measures.

A third issue in these results is the apparent lack of relationship among the other two perception categories (powerful others and unknown) and the performance assessment score. The literature on these perceptions indicates that the effect is not a clearly "two-tailed" effect, as it were. That is, it is not fair to say that one starts in the middle and as one moves toward perceptions of internal control, one's performance is enhanced in permissive scenarios; and as one moves toward perceptions of external control, one's performance decreases. Rather, perceptions of internal control do enhance performance when they are allowed to have effect, but perceptions of external control do not cause *decrements* in performance. They simply prevent one from being creative or strategic when circumstances allow. Thus it is entirely reasonable to see negative, but non-

significant correlations among these perceptions of external control and the performance assessment score.

To summarize, these results provide empirical support for the claims that a performance assessment score captures more than objective scores do. In this case, one's perceptions of control are represented in the performance assessment score. That, however, raises more questions than it answers.

Questioning Construct-Relevant Variance

Messick's language of construct variance provides a very useful framework from which to view this particular situation and the performance assessment debate in general. At the surface, the debates about performance assessments appear to be about construct-relevant and construct-irrelevant variance. Proponents of performance assessments, like Wiggins previously cited, long for assessments which allow student choice, student learning styles, aptitudes and interests to come through. That makes a great many people uncomfortable, and they ask if these are relevant to the construct being measured. Two papers in particular make this argument.

Lu & Suen (1995), after showing that cognitive style differentially affects the scores on a performance assessment, warns that this could be a "potential concern for validity outcomes (p. 13)." Although they discuss the ramifications with appropriate caution and circumspection, these ramifications are almost all negative. The conclusion is that the presence of cognitive styles in performance assessment scores is not a good thing.

Parkes (1996) in discussing the same data as the present paper argued that there is a choice to be made. Perception of control affects performance assessment scores. Is the variance represented construct-relevant or construct-irrelevant? The question in that paper was: Who gets to decide?

These concerns only hold, however, when we assess the same construct in different formats. Any disagreement among formats, especially if it can be attributed to cognitive style or

perceptions of control, is problematic. And much of the rhetoric in the debate seems to be about the format of the assessments. Performance assessment gets the short end of this stick because it is new, "challenging" the established objective formats of assessment. Since they are accepted, any divergence between them and performance assessments is a strike against performance assessments.

If the focus were shifted, however, from format to content these issues dissipate. A re-reading of the proponents of performance assessments (e.g. Wiggins, 1989; Shepard, 1989; Meisels, Dorfman & Steele, 1995) reveals that they are not asking for the assessment of the same constructs in different formats. They are asking for the assessment of different constructs. Wiggins (1989) wants to see the performance of exemplary tasks. That is quite different than the desire to find out what our students know (Shepard, 1989).

In the context of this present study, the difference is between what Spanish vocabulary and verb forms students know and which verb forms and vocabulary than can use creatively and coherently. There is a big difference. The objective test in this study tapped the former while the paragraph the students wrote tapped the latter. The question now is not which format is more valid, but which construct is the one we really want. If knowledge of Spanish verb forms and vocabulary is the focal construct, then the objective measure is more valid because the performance assessment includes construct-irrelevant variance, to return to Messick's language. If students' ability to utilize the Spanish language to accomplish a rhetorical goal is the focal construct, then the objective test suffers from construct underrepresentation.

Conclusion

As more and more research is conducted on performance assessments, it is likely that more and more results like the ones presented in this study will emerge. Educational psychologists and cognitive psychologists are continuing to build networks of relationships between "pure" cognitive functioning and other variables, such as motivation, cognitive style, interest, etc. An assessment

event is a cognitive event, and thus it is reasonable to expect it to be susceptible to the same forces as other cognitive events.

This will continue to occur because educational assessments are beginning to move away from assessing inert, static, declarative knowledge. Educational assessment is continually and increasingly challenged to show what students can *do* not simply what students *know*. The shift in focus is going to require shifts in method, but the attention needs to be placed shrewdly on the focus and less on the method.

References

Boggiano, A. K. and Katz, P. (1991). Maladaptive achievement patterns in students: The role of teachers' controlling strategies. *Journal of Social Issues, 47 (4)*, 35-51.

Bond, L. (1995). Unintended consequences of performance assessment: Issues of bias and fairness. *Educational Measurement: Issues and Practice, 14 (4)*, 21-24.

Connell, J. P. (1985). A new multidimensional measure of children's perceptions of control. *Child Development, 56*, 1018 - 1041.

Csikszentmihalyi, M. and Nakamura, J. (1989). The dynamics of intrinsic motivation: A study of adolescents. In C. Ames and R. Ames (Eds.), *Research on motivation in education: Goals and cognitions* (Vol. 3, pp. 45-71). New York: Academic Press.

Deci, E. L. and Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53 (6)*, 1024 - 1037.

Lepper, M. R. and Hoddell, M. (1989). Intrinsic motivation in the classroom. In C. Ames and R. Ames (Eds.), *Research on motivation in education: Goals and cognitions* (Vol. 3, pp. 73-105). New York: Academic Press.

Lu, C. and Suen, H. K. (1995). Assessment approaches and cognitive styles. *Journal of Educational Measurement, 32 (1)*, 1 - 17.

Meisels, S. J., Dorfman, A., & Steele, D. (1995). Equity and excellence in group-administered and performance-based assessments. In M. T. Nettles & A. L. Nettles (Eds.), *Equity and excellence in educational testing and assessment* (pp. 243-261). Boston: Kluwer.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50 (9),* 741-749.

Parkes, J. (1996, October). *Performance assessment and student motivation: A question of construct relevant variance*. Paper presented at the Annual Meeting of the Northeastern Educational Research Association, Ellenville, NY.

Shepard, L. (1989). Why we need better assessments. *Educational Leadership,* , 4-9.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi delta Kappan, 70 (9)* , 703-713.

Appendix A

Performance Assessment Task

Before doing the performance assessment , students performed the following exercise:

Using the map, choose two of the situations described below and create a dialogue of at least two speeches for each group member based on one of the following situations. Your dialogue should include as much of the vocabulary and the verbs in this chapter as is practical.

1. Your family is on vacation. Plan a day of sightseeing in which you will visit three places. Write down how to get to them from your hotel. Decide what time you will leave, when to return, etc.

2. You and your cousin are lost. Ask the policeman for directions to where you want to go and where to catch a bus or subway.

3. Your aunt and uncle have arrived at the airport. Tell them how to get to the place where you will meet them. When you find them, discuss the weather, what you can do tomorrow, and when they return home.

4. Your grandma is visiting. Decide on an itinerary for the next day's activities. (What you can do, when to leave, return, where the places are, etc.) Finally, tell her where she will sleep during her visit and say goodnight.

After that in-class group activity, students were asked to write either a diary entry or a letter to a friend describing the day's activities.

Appendix B

Scoring Rubric

Verb Forms	1	2	3	4	5
Vocabulary	1	2	3	4	5
Coherence	1	2	3	4	5
Creativity	1	2	3	4	5